DOCUMENT RESUME

ED 464 131 TM 033 814

AUTHOR	Vendlinski, Terry
TITLE	A Markov Model Analysis of Problem-Solving Progress.
PUB DATE	2002-04-01
NOTE	16p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 1-5, 2002).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Chemistry; *High School Students; High Schools; Learning Strategies; *Markov Processes; *Mathematical Models; Multimedia Materials; *Problem Solving
IDENTIFIERS	Neural Networks

ABSTRACT

This study used a computerized simulation and problem-solving tool along with artificial neural networks (ANN) as pattern recognizers to identify the common types of strategies high school and college undergraduate chemistry students would use to solve qualitative chemistry problems. Participants were 134 high school chemistry students who used the Interactive Multi-Media Exercises (IMMEX) software with a problem set related to identification of hazardous materials. Then, based on the calculated probabilities that students would transition between these strategy types over time, hidden chain Markov analysis allowed two objectives to be accomplished. A model of the capacity of the current curriculum to produce students able to apply chemistry content to real world problems was developed, and then this model was used to suggest pedagogical interventions that might be most effective at promoting better student understanding and as a metric by which to evaluate the potential effectiveness of these and other interventions. (Contains 9 tables and 15 references.) (SLD)

A Markov Model Analysis of Problem-solving Progress

Terry Vendlinski

This paper is prepared for the:
Annual Meeting of the American Educational Research Association in New Orleans, LA
April 2002

A Markov model analysis of problem-solving progress and transfer

Terry Vendlinski
UCLA / IMMEX Lab
5601 W. Slauson Avenue #255, Culver City, CA 90230

Abstract: Valid formative assessment is essential to the advancement of student learning and the development of pedagogical content knowledge in teachers (Nathan, Koedinger et al. 2001). Most currently accepted pencil-and-paper standardized tests, however, are not designed as formative assessment tools (Bransford, Brown et al. 1999; AERA 2000), and many performance-based assessments suffer from validity (Barton 1999), pedagogical (Lowyck and Poysa 2001), logistic, time and cost problems (Quellmalz, Schank et al. 1999). Moreover, recent thinking in the field of educational assessment suggests that formative assessments must focus less on how closely student responses match a pre-determined model and more on the competency of the performance as a whole (Pellegrino, Chudowsky et al. 2001). While the unstructured nature of such responses makes the evaluation of these types of performances difficult, the need for such evaluations is likely to increase. As computer hardware become cheaper, connectivity easier, and software development more rapid, computerized learning and assessment simulations are likely to become more ubiquitous throughout the American educational system. Consequently, methodologies to analyze and exploit the rich source of data from such performances must be devised with an eye to informing pedagogical interventions in a timely and valid manner.

This study used one such computerized simulation and problem-solving tool along with artificial neural networks (ANN) as pattern recognizers to identify the common types of strategies high school and college undergraduate chemistry students used to solve qualitative chemistry problems. Then, based on the calculated probabilities that students would transition between these strategy types over time, hidden chain Markov analysis allowed us to accomplish two objectives. First, we developed a model of the capacity of the current curriculum to produce students able to apply chemistry content to real world problems. Second, we used the model to both suggest the pedagogical interventions that might be most effective at promoting better student understanding and also as a metric by which to evaluate the potential effectiveness of these and other interventions.

Keywords: formative assessment, science education, markov modeling, artificial neural networks.

Background

The Interactive Multi-media Exercises (IMMEX™) software is a problem presentation, learning, and assessment tool that allows teachers to develop and present domain specific problems to their students. Because IMMEX software allows teachers to author their own simulations, the content of each IMMEX problem-set can be tailored to meet specific curricular objectives and to take account of specific classroom and student contextual variables an educator feels are important. For example, teachers may delete certain reference items for Advanced Placement (AP) students because students are required to memorize these items for the AP test. In addition, the IMMEX environment also makes the development of different versions (cases) of each problem easy and allows students multiple opportunities to apply their knowledge in similar but not identical situations or to attempt problems with different degrees of difficulty. Generally, the number of informational items available to the student when solving cases in a particular problem-set does not change from case to case, but the number of available items is often quite different between problem-sets. While standard Item Response analysis has produced good models of case difficulty both in the problem set discussed here as well as other

problem-sets, most teachers currently choose the problem-sets and the specific cases their students will attempt to solve more subjectively.

Students solve IMMEX cases by formulating hypothetical answers, accessing as much information as they feel necessary to test such answers, and then selecting an answer from a list of possible answers or by typing in their solution. As students proceed through IMMEX cases, the software records a student's every step as s)he attempts to solve each case. This feature allows for both real-time and off-line analysis of how students solve a particular case, as well as how student ability changes over time. Since students access IMMEX problems using the World Wide Web, the IMMEX database contains thousands of student performances on each of hundreds of problem-sets in different knowledge domains. For this study, the teacher developed 23 cases of an IMMEX qualitative chemistry problem-set as a tool to access how well her students could apply the concepts taught in first year, high school chemistry.

Methods and Data Sources

One hundred thirty-four first year chemistry students at a suburban Southern California high school were tasked to identify various unknown chemical compounds using the IMMEX computer simulation software. Student grade point averages, first semester grades, standardized test scores and student demographic data suggest this population of students is typical of student populations at suburban American high schools with the exception that African American students were under represented and Asian American students were overrepresented in this group; however, we have found trends similar to those reported here in student groups where African American students were overrepresented and Asian American students were underrepresented in the study cohort.

The IMMEX problem-set used by the students in this study is called *Hazmat* (short for Hazardous Materials). *Hazmat* is a qualitative chemistry problem-set in which students are told that there has been an earthquake that has caused a number of chemicals, some of which may be hazardous, to fall off stockroom shelves. As the labels have been obliterated or mixed up with other compounds, water is beginning to flood the storeroom and time is of the essence, the school has asked for student help in identifying the spilled chemicals. In addition to general stockroom inventory information, there are three physical tests and eight chemical tests the students can conduct on the unknown substance. The students may also review any of eight general reference items in their attempt to identify the unknown and students must identify the correct unknown from among 57 possibilities on one of two tries at a solution. After the presentation of the problem, students may proceed through the problem space in any manner they choose before ultimately proposing the identity of the unknown. The number of menu items is consistent across all 23 cases in the *Hazmat* problem-set. In this study, the students' teacher decided to use only those *Hazmat* cases that produced positive results when students chose to conduct a flame test. Both IRT analysis and the teachers experience suggest these are the easiest of the *Hazmat* cases for this group of students to solve. Each student received individual cases in random order.

While a single student action is occasionally informative in IMMEX problem-solving (such as when a student chooses to solve the problem as an initial move and without viewing any information), experience suggests that the sequence of actions or the presence of a group of individual actions are usually much more telling. The number of possible information items in a problem-set and the degree of difficulty of the cases a student must solve are generally good indicators of the number of menu items students will choose to view before solving individual cases of a problem-set. The students in this study typically chose to view 17 items before attempting a solution.

Experience also suggests that student performances on cases of IMMEX problem-sets are seldom entirely random in nature. In fact, while students may eventually look at all the information contained in a problem space, they will often view menu items sequentially rather than follow a more haphazard strategy. Nevertheless, given the number of pieces of information available to the student, the possible number of unique performances is factorially large. This number is even larger when considering not just the items a student chose, but the *order* in which the student chose them. The overwhelming nature of such a task becomes especially apparent when the number of student performances increases beyond one hundred or so. Consequently, as problem-spaces become large and student performances multiply, some method must be used to help discern patterns in such data. Nevertheless, cognitive scientists like Fischer and Bidell (1997) suggest one must develop ways to analyze the patterns of stability and order within the variation of human activity in order to simplify and understand that activity without discarding behavioral complexity. However, the literature suggests that it is unwise to merely compare students or other novices to their teacher or another expert.

Novices treat problem solving differently than experts (Messick 1989; Glaser and Silver 1994; Baxter and Glaser 1997), and such comparisons are bound to limit the acceptable student approaches or to discard important complexities of student problem solving. Therefore, rather than develop a model of behavior and then fit subsequent student performances to that static model, we have chosen to use the demonstrated pattern recognition ability of artificial neural networks to identify groups of similar performances in the data (Principe, Euliano et al. 2000). As described elsewhere (Stevens, Lopo et al. 1996; Vendlinski and Stevens 2000; Vendlinski 2001), the networks are able to cluster the same performances together almost 90% of the time and researchers can easily identify those aspect of the performance clusters which make the performances similar.

These descriptive aspects of student performances are useful both to validate the clusters and as a general descriptive of the clustered performances. *Mock performance analysis* was used to validate the meaning we ascribed to each cluster of student performances. In this type of analysis, a mock performance is created to represent each of the different clusters. Each of these mock performances is then, in turn, fed back through the appropriately trained artificial neural network. Repeatedly, we have found that if more than 60% of the performances in a cluster chose a particular menu item as part of their strategy, that item will be an important descriptor of the performances in that cluster. Using this technique, researchers have found that 100% of the mock performances cluster where anticipated, if that cluster contained at least three performances. Not only does this

technique validate the interpretation of each cluster, but, by using the important menu items to describe each cluster, the differences between clusters become readily apparent. Moreover, the sensitivity of the clustering network to variations in student performances can be quantified by adding or subtracting menu items in a mock performance. We call each cluster descriptive a *strategy*.

By assigning each student performance to a cluster, and then ordering these performances chronologically, longitudinal models of student problem-solving emerge. In this study, a student's first performance strategy was compared to the most common strategy used by the student to solve all the *Hazmat* cases the student attempted. When viewed individually, this type of analysis addresses the progress individual students make over time. When analyzed as a group, it describes students more generally, and becomes an indicative of class progress or, with multiple classes, teachers and schools, of more generalized learning trends. With these larger student groups, the likelihood students will transition from using one strategy (the beginning state) to subsequently using the same or another strategy is easy to calculate. When represented in condensed form these likelihoods form a transition matrix.

By using the transition matrix and elementary matrix multiplication, the distribution of students across clusters can be calculated. Moreover, with a large enough sample both latent trait theory and our prior research suggest it is plausible to make the assumption that without external intervention, the transition likelihoods for a group of students remain constant over time. It then becomes possible to apply Hidden Markov Chain analysis to determine the distribution of students after each student works a number of successive cases in a problem-set. Obviously, since multiple paths are usually possible, the exact path a student follows to arrive at a particular cluster is not considered (hence they are *hidden* Markov chains). Nevertheless, as students continue to work new cases, the distribution of a group of student performances most often achieves a *steady state*. At this point, unless the transition matrix changes, the overall distribution of student performances within a group will no longer change between clusters, so an evaluation of the group is possible. In other words, steady state diagrams provide a metric by which we can evaluate the performances used by a teacher's students. Although not a major focus of this paper, the diagrams would allow comparisons of different classrooms to one another even though different teachers taught and students used different means to solve the cases in a problem-set. Moreover, these diagrams allow us to model not only the predicted effect of proposed pedagogical interventions, but also how such an intervention might ultimately affect the strategies used by these students. This technique is discussed at the end of the next section.

Discussion

While the student strategies represented by each cluster can vary widely between clusters, the strategies are often closely related and may only differ by one or two items of information. In fact, in some recent performances, the major difference between two strategies was whether a student viewed the problem summary (epilog). Another major difference between strategies is the success students have solving a case using that

particular strategy. For this research, we determined the effectiveness of student strategies by calculating the odds a particular strategy would produce the correct answer to *Hazmat*. Good strategies tended to produce a correct answer to the case, while poor strategies did not. We use odds here to allow the comparison of solve rates between different types of cases (as the cases were delivered randomly, not all cases were delivered with the same frequency), and because the natural logarithm of the odds equals student ability (θ) in the one-parameter logistic model. Depending on the variability of cases within a problem-set, the strategies students use to solve each case and the success rate of each strategy may differ. In some IMMEX problem-sets, a single, well developed strategy will allow a student to solve all the cases within a problem set. In other problem-sets, students will have to adapt their strategies to account for changes in the cases. The *Hazmat* problem-set is an example of the latter. For example, a student may be able to identify the unknown compound Potassium Hydroxide using a flame test and Litmus paper, but would have to modify this strategy to identify Potassium Nitrate. Nevertheless, one might expect that a student who developed the ability to solve one case would demonstrate the ability to effectively solve other, different cases, especially if the student really understood the concepts required to solve these types of problems.

As expected, in IMMEX problem-sets where the cases require students to modify their strategy as the unknown changes, students often do not duplicate a specific strategy verbatim; rather the students adapt their strategy to the case they are trying to solve. Nevertheless, there are enough similarities between the different strategies that suggest a more general classification might be appropriate. In particular, a number of students use strategies that investigate very few items of information before an attempt is made to solve the case. In fact, none of these so called *limited* types of strategies investigate enough information to conclusively solve the problem. At the other extreme, students investigate more than enough information to solve the problem and often continue to view items after they have sufficient information to reach a definitive answer. We have termed these strategies *prolific*. Students using either limited or prolific types of strategies are unlikely to solve the case being attempted. In other words, overall they had less than even odds of correctly identifying the unknown. However, when students used strategies that focus only on key pieces of information, they were more likely than not to solve every case they attempted. These strategies were classified as *efficient* types of strategies. Efficient strategies tend to be both case and contextually sensitive. For example, one group of AP chemistry students was very successful in determining the presence of unknown bases by using a strategy which involved the addition of an acid to the unknown, whereas another group of first year chemistry students at the same school were more successful using a strategy that involved using Litmus paper to identify the presence of a base. While both strategies were effective for both groups of students and were very focused, the specific information underlying the success of each strategy differed. Anecdotally, the AP teacher indicated she had stressed the addition of acid, whereas the teacher of first-year students had focused on the use of Litmus paper. It is this characteristic of strategy types that makes comparisons between classes as well as comparisons within a single class possible. The next section illustrates such a within class comparison for use as a formative tool.

Initially, approximately one-third of the students studied here chose to use limited types of strategies, one-third chose prolific strategies and the final third, efficient strategies. As expected, a significant number of students who solve the first *Hazmat* case presented to them using an efficient type of strategy, use efficient types of strategies to solve most of the other *Hazmat* cases presented to them. More surprisingly, students using limited or prolific types of strategies will predominately continue to use limited or prolific strategy types, respectively, in their attempts to solve future cases *even though those strategies seldom produce a correct answer*. These relationships are shown in Table 1 and, as indicated there, are significant.

First Strategy: \ Most Frequent Strategy:	Limited	Efficient	Prolific
Limited	30	6	8
Efficient	5	23	9
Prolific	5	8	35

Table 1. This table shows the relationship between the type of strategy a student used to solve his or her first *Hazmat* case and the strategy the same student use most often (the mode) to solve subsequent *Hazmat* cases. Overwhelmingly, without pedagogical intervention, the strategy type used by a student in her or his initial performance predicts the strategy type that student will continue to use most often on subsequent *Hazmat* cases. The chi-square statistic suggests is relationship is not random. ($\chi^2 = 70.5$; d.f. = 4; $p < .001$).

This same trend is evident in other classes, at other ability levels and on different IMMEX problem sets. For example, Table 2 shows the relationship between the strategy type first year undergraduate chemistry students used to solve their first case of a more complex qualitative chemistry IMMEX problem-set and the strategy type used most often by these same students to solve subsequent cases of this problem-set.

First Strategy: \ Most Frequent Strategy:	Limited	Efficient	Prolific
Limited	28	3	7
Efficient	2	4	1
Prolific	11	1	28

Table 2. This table shows the relationship between the type of strategy a first-year chemistry college undergraduate used to solve his or her first qualitative chemistry case and the strategy the same student used most often (the mode) to solve cases of the same problem-set. Overwhelmingly, the strategy type used by a student in her or his initial performance correlates with the strategy type that student will continue to use on subsequent cases.

The degree of difficulty of the cases or problem-set also seems to have an effect on the strategy type students use to solve the cases of IMMEX problem-sets. Research suggests that as a problem space becomes easier for students to manage (either because the ability level of the students increases or the problem space become less complex), students are less likely to use limited types of strategies. In fact when the college undergraduates just described attempted to solve *Hazmat* cases, no student used a limited type of strategy to solve the problem. On the other hand, when the first-year high school chemistry students in this study were asked to solve the more complex qualitative chemistry problem mentioned above, they used almost entirely limited types of strategies to do so. Table 3 shows this

First Strategy: \ Most Frequent Strategy:	Limited	Efficient	Prolific
Limited	56	11	2
Efficient	10	14	3
Prolific	6	0	12

Table 3. The same relationship between the type of strategy a high school student used to solve the initial case of a more complex IMMEX problem and the type of strategy s)he uses most often is evident in this table. As the problem-space has become more complex, students begin to favor the use of limited types of strategies.

($\chi^2 = 64.4$; d.f. = 4; $p < .001$).

The types of strategies used by a group of students, therefore, seems to suggest whether cases of a particular problem-set like *Hazmat* are too easy or too difficult. When almost

all students use limited types of strategies to solve presented cases (a type of "floor effect"), the cases or the entire problem-set is probably too difficult. Likewise, when most students use efficient or prolific types of strategies to solve cases in a problem-set (a type of "ceiling effect"), those cases are probably not challenging enough for the students or indicate that the students have mastered the material. When, however, student performances demonstrate a range of strategy types, one may generally conclude a problem-set or a group of cases is appropriate for that student group. Although such considerations may be less important for summative assessment, they are critical when using these results to formulate curricular interventions.

In fact, when the performances of a group of students represent diverse solution strategies, Markov hidden chain analysis has proven to be an effective tool for evaluating the pedagogical interventions suggested to improve student performances. For example, Table 4 shows the probability that the first year high school chemistry students in this study will start off with a particular type of strategy, then use another or repeatedly use the same type of strategy on a subsequent performance.

To: / From:	Start	Limited	Efficient	Prolific
Start	.00	.34	.29	.37
Limited	.00	.56	.19	.25
Efficient	.00	.20	.54	.26
Prolific	.00	.17	.26	.57

Table 4. The probability that a student in this group will transition from the strategy type given in the left column to the strategy type identified at the top of each column is given where that row and column intersect. As students cannot return to the "Start" state once they solve their first case, the probability that they move from any state back to the "Start" state is 0%. Note, however, that for this group, the likelihood a student will use the same type of strategy to solve the very next problem (the diagonal cells) is more than 50% for each of the strategy types.

Assuming, as the research suggests, that these probabilities remain consistent over time, one can use such a "transition matrix" to predict how many students will use a particular type of strategy on their next case by multiplying the vector representing the number of students using each strategy type on their present case by the transition matrix. As we have found that even passive interventions leave this matrix unchanged over time, we believe this assumption to be reasonable. Multiplying the resulting vector by the transition matrix again will yield a prediction of the number of students using each strategy type after two more cases, and so on. Although one cannot use this method to trace how an individual student came to arrive at a particular strategy type (hence the name "hidden" chain analysis), repeatedly multiplying by the fixed transition matrix generally produces a *steady state*.

To: / From:	Start	Limited	Efficient	Prolific
Start	.00	.30	.33	.37
Limited	.00	.30	.33	.37
Efficient	.00	.30	.33	.37
Prolific	.00	.30	.33	.37

Table 5. The probability that a student in this group will transition from the strategy type given in the left column to the strategy type identified at the top of each column after attempting to solve seven cases of *Hazmat* is given where that row and column intersect. Assuming that the probability a student will transition from a given state to a subsequent state remains constant during these seven cases (Table 4), the above steady state is reached. The analysis of this group of students suggests that 30% of the students will settle into using Limited types of strategies, 33% will use efficient strategies, and 37% will use prolific strategies when solving *Hazmat* cases.

As shown in Table 5, because students tend to adopt a single type of solution strategy from the outset of solving *Hazmat* cases, approximately one third of the students will eventually settle into each of the three strategy types. While such steady states allow comparisons between classrooms or teachers, a more formative use is also suggested.

When a teacher analyzes how the strategies identified by artificial neural analysis differ, and combines the insight of that analysis with the Markov technique just demonstrated, various pedagogical interventions can be modeled and the predicted effectiveness of each intervention compared. For example, when reviewing the student performances that generated the data in Table 4, it becomes obvious that one of the efficient strategies students use to identify part of the unknown is the use of red Litmus paper. Red Litmus turns blue in the presence of hydroxide, so a successful red Litmus test, along with an informative flame test, should allow the student to correctly identify these types of unknowns. Consequently, the teacher may decide to revisit the use and meaning of red Litmus in her curriculum. Markov hidden chain analysis allows us to model and predict the outcomes of such an intervention. If we assume, for the moment, that 90% of the students in this class developed an understanding of and could effectively use red Litmus after the teacher's intervention, this would imply a change in the types of strategies used by those students when trying to solve *Hazmat* cases that involve hydroxides. More specifically, because the red Litmus is now meaningful to the students currently using limited or prolific strategies, these students would modify their existing strategy so it became more efficient. This change in student behavior would produce the transition matrix in Table 6 and the steady state matrix in Table 7.

To: / From:	Start	Limited	Efficient	Prolific
Start	.00	.21	.47	.32
Limited	.00	.39	.41	.20
Efficient	.00	.24	.49	.27
Prolific	.00	.12	.44	.44

Table 6. The probability that a student in this group will transition from the strategy type given in the left column to the strategy type identified at the top of each column from their first to second case of *Hazmat* after instruction on red Litmus paper is given where that row and column intersect.

To: / From:	Start	Limited	Efficient	Prolific
Start	.00	.24	.46	.30
Limited	.00	.24	.46	.30
Efficient	.00	.24	.46	.30
Prolific	.00	.24	.46	.30

Table 7. The probability that a student in this group will transition from the strategy type given in the left column to the strategy type identified at the top of each column after instruction on red Litmus paper and after attempting to solve seven cases of *Hazmat* is given where that row and column intersect. Assuming that the probability a student will transition from a given state to a subsequent state (Table 6) remains constant during these seven cases, the above steady state is reached. The analysis of this group of students suggests that almost half the students will use efficient strategies when solving *Hazmat* cases after this specific intervention.

Other similar calculations could be made by reducing the percentage of students who benefit from the instruction on red Litmus paper. Moreover, the effects of other interventions can be modeled and compared with re-teaching this topic. For example, students in this data set also use the strategy of adding acids to the unknown, and we have also found this to be a common strategy among high school Advanced Placement students. As most chemistry teachers would know, this strategy is particularly effective because when an acid is added to a compound composed, in part, of hydroxide, it will become hot. Moreover, the same addition will cause a compound containing a carbonate to bubble. If the teacher of this group of students was to teach these students this concept and 90% of the students effectively applied the new strategy on subsequent *Hazmat* cases, the new transition matrix shown in Table 8 would result. Assuming that these transitions remain constant, the steady state resulting from this new transition matrix is shown in Table 9.

To: / From:	Start	Limited	Efficient	Prolific
Start	.00	.19	.5	.31
Limited	.00	.36	.5	.14
Efficient	.00	.22	.53	.25
Prolific	.00	.10	.48	.42

Table 8. The probability that a student in this group will transition from the strategy type given in the left column to the strategy type identified at the top of each column from their first to second case of *Hazmat* after instruction on acid reactions is given where that row and column intersect.

To: / From:	Start	Limited	Efficient	Prolific
Start	.00	.22	.51	.27
Limited	.00	.22	.51	.27
Efficient	.00	.22	.51	.27
Prolific	.00	.22	.51	.27

Table 9. The probability that a student in this group will transition from the strategy type given in the left column to the strategy type identified at the top of each column after instruction on acid reactions and after students attempt to solve seven cases of *Hazmat* is given where that row and column intersect. Assuming that the probability a student will transition from a given state to a subsequent state (Table 8) remains constant during these seven cases (Table 8), the above steady state is reached. The analysis of this group of students suggests that almost half the students will use efficient strategies when solving *Hazmat* cases after this specific intervention.

Here again, effectiveness rates lower than 90% could be applied, and different effectiveness rates could be applied to each of the two interventions based on empirical or a teacher's anecdotal evidence about the ease students learn and apply each of the interventions. Conversely, the difference between predicted and actual outcomes could be used as a metric for the degree of difficulty specific students or students in general have learning or applying various concepts.

The steady state tables resulting from this analysis provide at least three key insights. First, comparing steady state tables suggest which intervention might produce the greatest amount of positive change. Among this group of students, those using efficient strategy types were more likely than not to correctly apply qualitative chemistry concepts to correctly identify unknowns. Students using limited and prolific strategies were more likely to misidentify unknowns. The two steady state tables above (Tables 7 and 9) suggest that teaching these students about the interaction of red Litmus and solutions containing hydroxide would result in slightly less than half the students adopting more efficient strategies to solve these types of *Hazmat* cases (Table 7) if 90% of the students effectively applied that instruction in problem-solving. On the other hand, if the student's teacher achieved the same degree of effectiveness in teaching her students how acids react with carbonates and with hydroxides, we would expect more than 50% of the students to adopt efficient strategies (Table 9).

The second piece of information provided by this analysis is a quantitative measure of the difference between predicted and actual outcomes. While the first intervention is predicted to move almost 50% of the students to adopt more effective strategies, what actually happens after such an intervention might be very different. Such discrepancies could provide new insights into more effective teaching and learning paradigms *based on the past problem-solving performances of the actual students being instructed.*

Finally, as can be seen in the variation between Tables 5, 7 and 9, these two interventions produce smaller changes in students using limited strategies than in those using prolific strategies. Because the cases were delivered entirely at random to the students, it is unlikely that the order in which a student received the cases will affect strategy type selection significantly before or after the teacher's intervention. Consequently, Markov analysis should allow us to predict what interventions might be most appropriate for those students using each of the various types of problem-solving strategies.

Conclusion

Quantitative and qualitative analysis suggest that the strategy types identified by artificial neural network analysis are both accurate and reliable. Moreover, this research suggests that such an analysis could function as a valuable formative tool by suggesting teaching interventions designed to benefit both individual students as well as larger, more diverse, groups of students. This study used adaptive artificial neural network analysis to identify the common strategies first year high school chemistry students used to solve qualitative chemistry problems. It demonstrated that the strategies used by these students were of three general types. Students adopting limited types of strategies did not have enough information to proffer a conclusive answer before doing so. On the other hand, students using prolific strategies had more than enough information to precisely identify the unknown. In both cases, however, students adopting either strategy type were unlikely to correctly identify the compound. Conversely, about one-third of the students in this study adopted very efficient strategies that allowed them to focus only on information that was pertinent to correctly identifying the unknown. Students adopting efficient strategies were more likely than not to identify the unknown substance.

This study also found that no matter which type of strategy the student used to solve *Hazmat* cases, students would adopt that strategy type beginning with the first case and they would continue to use similar strategies on subsequent cases. This same trend has also been documented in other high school science domains, and among chemistry students of varying abilities (e.g. high school Advanced Placement, community college, and first year undergraduates). Without pedagogical intervention, students appear highly unlikely to change problem-solving strategies, *even if those strategies seldom produce a correct answer*. Nevertheless, this research suggests that the analysis of strategies combined with Markov hidden chain analysis could function as a valuable formative tool.

When combined with a teacher's insight of how the strategy types of students differ, Markov analysis not only suggests which interventions might be most effective for students, but also provides a metric that allows us to compare the potential effectiveness

of each intervention. In this case, the use of red Litmus and the addition of acid to an unknown were two strategies repeatedly used by successful students. Consequently, we chose to evaluate how re-teaching the concepts underlying these strategies (not just the strategies themselves) would improve the problem-solving of the students. Markov analysis suggests that both interventions are likely to improve student problem-solving, but that teaching students the effects of adding an acid to an unknown is likely to result in more of the students using better strategies.

Although time and curriculum limitations prevented the interventions described above from being implemented, determining how closely the steady states predicted by Markov analysis match reality is both necessary and planned. More importantly, research must demonstrate how closely the effects of the suggested pedagogical interventions match those predicted by the calculated steady states or if the relative differences between interventions approximates the differences in student problem-solving behaviors actually seen in chemistry classrooms. In the meantime, the methodology proposed here offers investigators in the field and educators in the classroom a metric that allows each to develop and to begin evaluating the effectiveness of such pedagogical interventions.

References

AERA (2000). AERA position statement concerning high-stakes testing in PreK - 12 education. Washington DC, American Education Research Association.

Barton, P. E. (1999). Too much testing of the wrong kind; too little of the right kind in K-12 education. Princeton, NJ, Educational Testing Service.

Baxter, G. P. and R. Glaser (1997). An approach to analyzing the cognitive complexity of science performance assessments. Los Angeles, CRESST/UCLA.

Bransford, J. D., A. L. Brown, et al., Eds. (1999). How people learn: brain, mind, experience, and school. Washington, DC., National Academy Press.

Fischer, K. W. and T. R. Bidell (1997). Dynamic development of psychological structures in action and thought. The Handbook of child psychology: theoretical models of human development. Lerner. New York, Wiley. **1:** 467 - 561.

Glaser, R. and E. Silver (1994). Assessment, testing and instruction: retrospect and prospect. Los Angeles, UCLA / CRESST.

Lowyck, J. and J. Poysa (2001). "Design of collaborative learning environments." Computers in human behavior **17**: 507-516.

Messick, S. (1989). Validity. Educational Measurement. R. L. Linn. New York, Macmillan Publishing Co.: 13 - 103.

Nathan, M., K. Koedinger, et al. (2001). The expert blindspot: when content knowledge and pedagogical content knowledge collide. 2001 AERA Annual Meeting, Seattle WA.

Pellegrino, J., N. Chudowsky, et al., Eds. (2001). Knowing What Students Know: The Science and Design of Educational Assessment. Washington DC, National Academy Press.

Principe, J. C., N. R. Euliano, et al. (2000). Neural and Adaptive Systems. New York, John Wiley & Sons.

Quellmalz, E., P. Schank, et al. (1999). "Performance Assessment Links in Science." Practical Assessment, Research & Evaluation **6**(10).

Stevens, R. H., A. C. Lopo, et al. (1996). "Artificial neural networks can distinguish novice from expert strategies during complex problem solving." Journal of American Informatics Association **3**: 13 - 38.

Vendlinski, T. and R. Stevens (2000). The use of artificial neural nets (ANN) to help evaluate student problem solving strategies. International Conference of the Learning Sciences, University of Michigan, Lawrence Erlbaum Associates.

Vendlinski, T. P. (2001). Affecting U.S. education through assessment: new tools to discover student understanding. Technology Management and Policy. Cambridge MA, Massachusetts Institute of Technology: 183.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM033814

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: A Markov analysis of problem-solving progress and transfer	
Author(s): Terry Vendlinski	
Corporate Source: UCLA / IMMEX	Publication Date: April 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 [X]	Level 2A [ ]	Level 2B [ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, please

Signature:	Printed Name/Position/Title: Terry Vendlinski Research Scientist	
Organization/Address:	Telephone: 310-649-6568	FAX: 310-649-6591
	E-Mail Address: Vendlins@ucla.edu	Date: 4/1/02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)